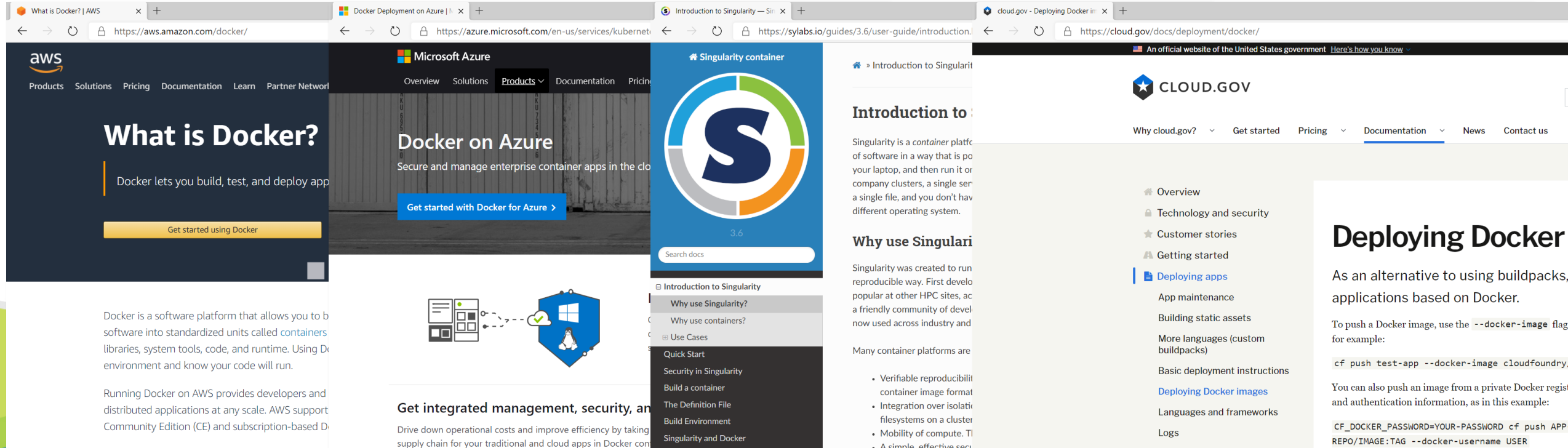
EPA DMAP DevOps*

USGS CI/CD CoP
August 4, 2020

Containerization As A Core Strategy

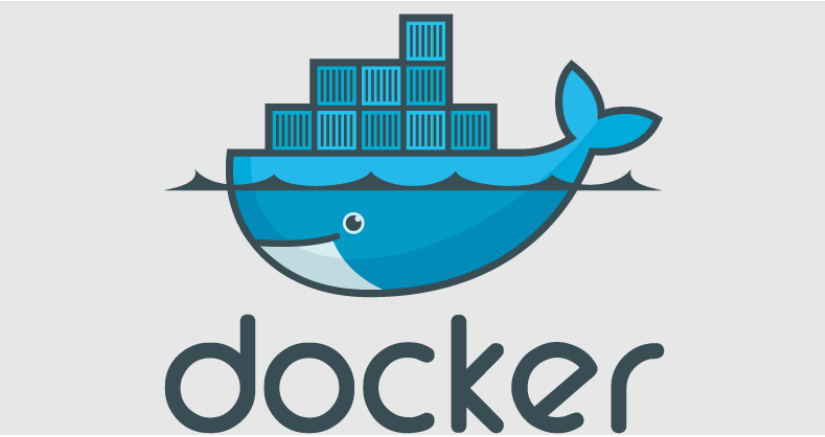
- EPA has a multi-cloud environment: AWS, Azure and Cloud.gov, all of which support Docker containers.
- EPA can also support Docker on Linux and OpenShift as well as in the High Performance Computing (HPC) environment via Singularity.



What is Docker?

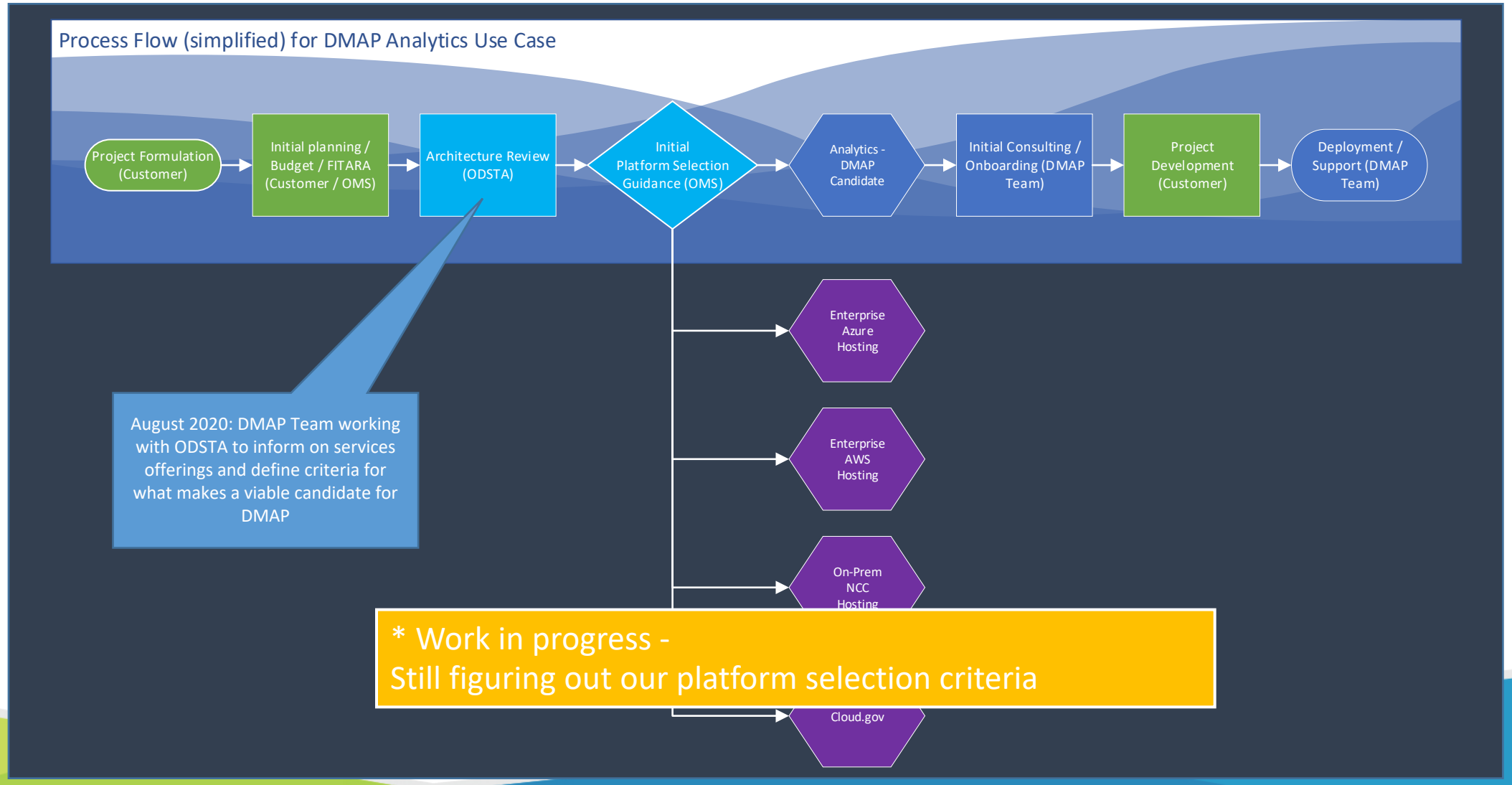
- Docker is often thought of as a Virtual Machine…
  …but it isn't.
- Docker is
  - a.) a packaging specification and
  - b.) a runtime for executing the packages (containers)
- Docker is much leaner than a VM, requires much less overhead
- Dockerfiles are similar to linux bash scripts, and are the recipe to build the executable Docker images (and dockerfiles are more easily sharable and customizable than VM images)
- Docker images are run on a Docker runtime. The Docker runtime can run multiple containers, and there are frameworks for scaling orchestrating containers across clusters of servers

Why use Docker?



- Reusability
- Lightweight, smaller footprint than a VM
- Support different application stacks
- “Deploy and run anywhere” - Run on different environments and managed services (AWS, Azure, etc)
- Can develop locally and deploy to cloud or other target environments
- For analytics, can use Docker to decompose functionality into microservices
- Can use Docker to spin up and replicate multiple instances to run in parallel

What Goes Where?

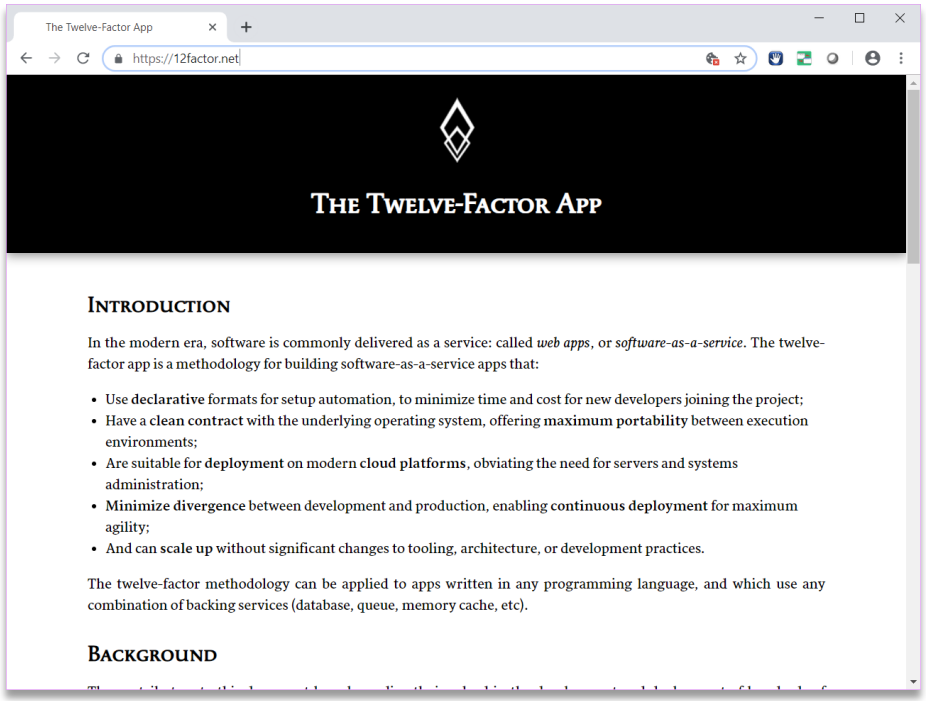


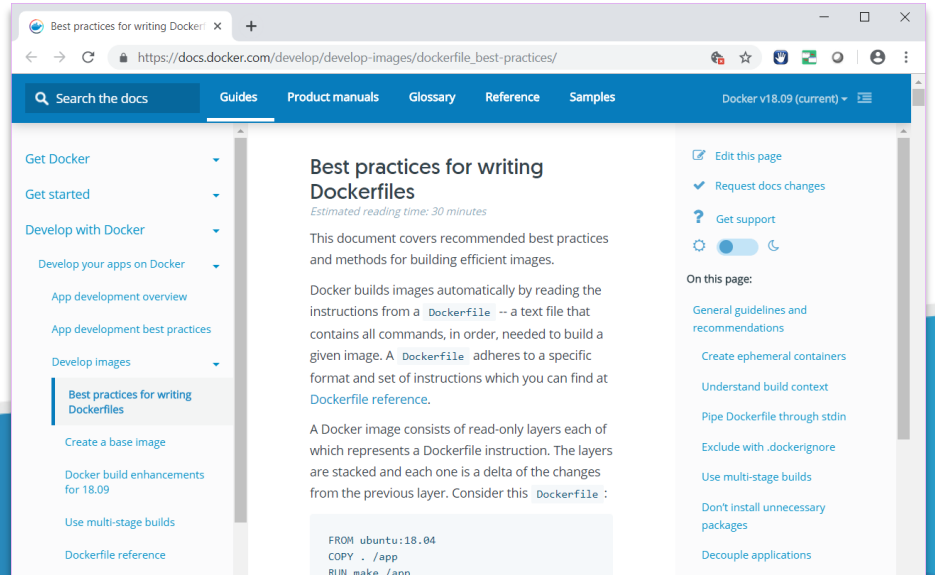
Containerization vs. Other Options

	Virtual Machine	Container	Lambda
Approach	Server	Swarm Cluster	Serverless
Team	Ops	DevOps	Dev
Architecture	N-Tier	Microservice	Event Driven
Pros	Consolidation	Portability	Scalability
Cons	Heaviness	Simplexity	Tight AWS Integration Short Execution Time

Containerization And Serverless Can Drive Best Practices

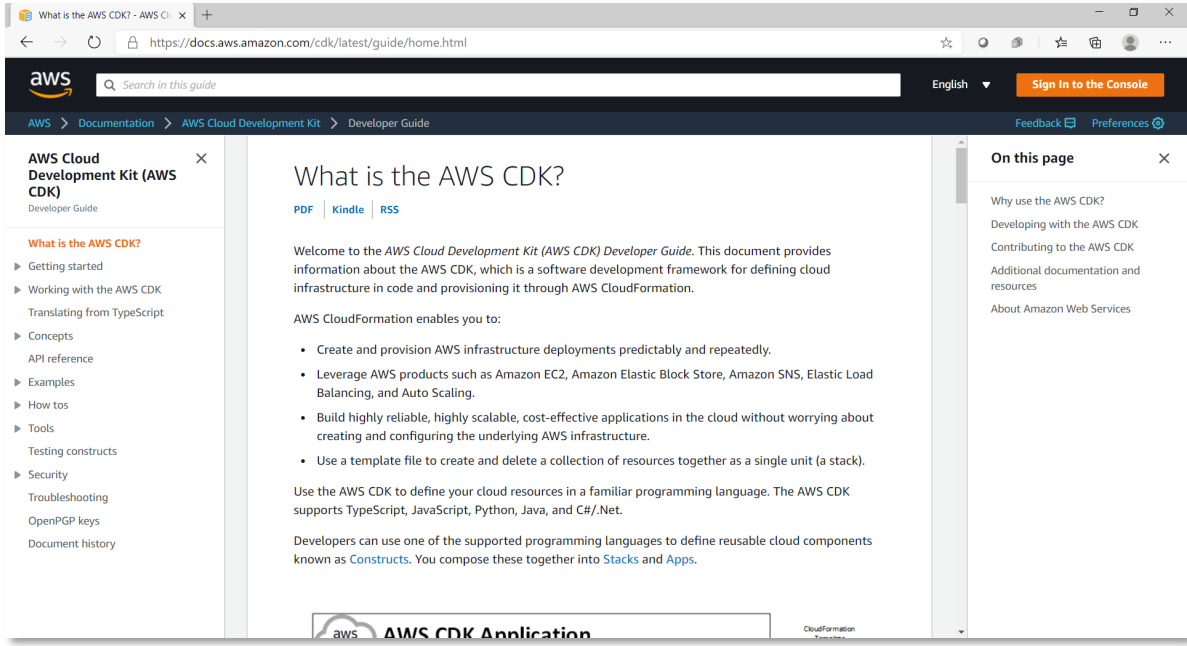
- Consider apps as ephemeral
- Cloud-native: Twelve-Factor Design
  - https://12factor.net
  - Store config in the environment
  - Treat backing services as attached resources
  - Code for fast startup and graceful shutdown
- Dockerfile Best Practices
  - https://docs.docker.com/
  - Minimize Layers
  - Build in stages
  - Decouple applications
- Many other guides exist
- In AWS, leverage other managed services where feasible/appropriate

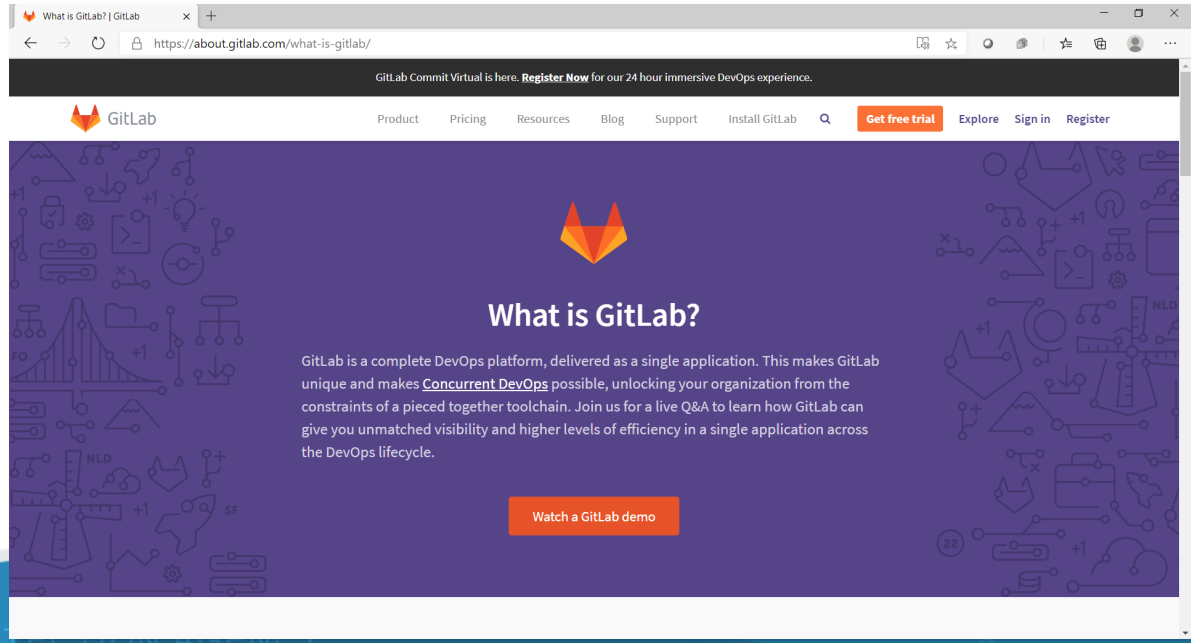




AWS DevOps

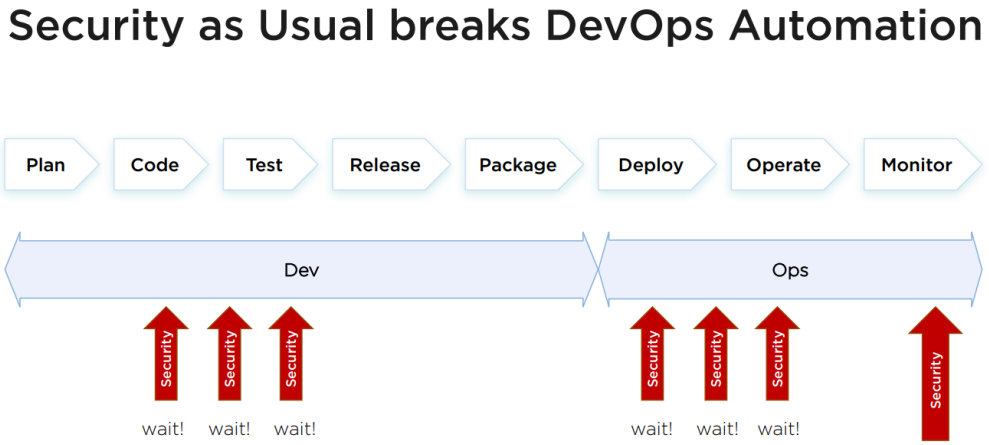
- We use AWS CDK to spin up stacks – manages environment and permissions, parameterized scripting for reproducibility
- We are using GitLab to manage our AWS Analytics build process
- Also using CodeCommit, S3 sync and other AWS tools
- National Computer Center Cloud DevOps team is also looking at Jenkins, Rancher, Atlassian and other tools
- How do we bring in diverse developer teams, contractors, et cetera?
- Works in progress…





How To Get To DevSecOps?

- Currently using Nessus to scan, Trend Micro Deep Security and other tools – build scans and nightly scans automated
- But what criteria for automating?
  - False positives in scan results?
  - Manual review and control gates for deploying to production?
  - Minor versus major release?
- Many things to still figure out…



Thank You

- Dave Smith
  - US Environmental Protection Agency
    - Office of Mission Support (OMS-EI)
    - Office of Information Management (OIM)
    - Information Access and Analytic Services Division (IAASD)
  - (202) 566-0797
  - smith.davidg@epa.gov